##### A Project report on

**Link Prediction in Social Media Network using Machine Learning approaches**

###### A Dissertation submitted to JNTU Hyderabad in partial fulfillment of the academic requirements for the award of the degree.

**Bachelor of Technology**

**in**

**Computer Science and Engineering**

Submitted by

L. Rahul Kumavath

(19H51A05L9)

1. Sreelatha

(19H51A05M3)

P. Sathvika

(19H51A05P5)

Under the esteemed guidance of

Mr. G. Saidulu

Assistant Professor



**Department of Computer Science and Engineering**

**CMR COLLEGE OF ENGINEERING & TECHNOLOGY**

(An Autonomous Institution under UGC & JNTUH, Approved by AICTE, Permanently Affiliated to JNTUH, Accredited by NBA.)

KANDLAKOYA, MEDCHAL ROAD, HYDERABAD - 501401.

#### 2019- 2023

**CMR COLLEGE OF ENGINEERING & TECHNOLOGY**

KANDLAKOYA, MEDCHAL ROAD, HYDERABAD – 501401

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**



#### CERTIFICATE

This is to certify that the Major Project Phase-1 report entitled **"Link Prediction in Social Media Network using Machine Learning Approaches"** being submitted by L. Rahul Kumavath 19H51A05L9, A. Sreelatha 19H51A05M3, P. Sathvika 19H51A05P5 in partial fulfillment for the award of **Bachelor of Technology in Computer Science and Engineering** is a record of bonafide work carried out his/her under my guidance and supervision.

###### The results embodies in this project report have not been submitted to any other University or Institute for the award of any Degree.

**Mr. G. Saidulu Dr. Siva Skandha Sanagala**

**Assistant Professor Associate Professor and HOD**

**Dept. of CSE Dept. of CSE**

#### ACKNOWLEDGEMENT

With great pleasure we want to take this opportunity to express my heartfelt gratitude to all the people who helped in making this project work a grand success.

We are grateful to **Mr. G. Saidulu, Assistant Professor** , Department of Computer Science and Engineering for his valuable technical suggestions and guidance during the execution of this project work.

We would like to thank **Dr. Siva Skandha Sanagala,** Head of the Department of Computer Science and Engineering, CMR College of Engineering and Technology, who is the major driving forces to complete my project work successfully.

We are very grateful to **Dr. Vijaya Kumar Koppula**, Dean-Academic, CMR College of Engineering and Technology, for his constant support and motivation in carrying out the project work successfully.

We are highly indebted to **Dr. V A Narayana,** Principal, CMR College of Engineering and Technology, for giving permission to carry out this project in a successful and fruitful way.

We would like to thank the Teaching & Non- teaching staff of Department of Computer Science and Engineering for their co-operation

We express our sincere thanks to **Mr. Ch. Gopal Reddy**, Secretary, CMR Group of Institutions, for his continuous care.

Finally, We extend thanks to our parents who stood behind us at different stages of this Project. We sincerely acknowledge and thank all those who gave support directly and indirectly in completion of this project work.

L. Rahul Kumavath 19H51A05L9

1. Sreelatha 19H51A05M3

P. Sathvika 19H51A05P5

Link Prediction in Social Media network using machine learning approaches

**TABLE OF CONTENTS**

**CHAPTER**

**NO. TITLE PAGE NO.**

LIST OF FIGURES ii

ABSTRACT iii

**1** **INTRODUCTION** 1

1.1 Problem Statement 1

1.2 Research Objective 2

1.3 Project Scope and Limitations 2

**2** **BACKGROUND WORK**

2.1. Variational Graph Auto-Encoder 3

2.1.1.Introduction 3

2.1.2.Merits,Demerits and Challenges 5

2.1.3.Implementation of Variational Graph Auto-Encoder 6

2.2. Adaptive Moment Estimation 9

2.2.1.Introduction 9

2.2.2.Merits,Demerits and Challenges 10

2.2.3.Implementation of Adaptive Moment Estimation 10

2.3. Node2Vec 13

2.3.1.Introduction 13

2.3.2.Merits,Demerits and Challenges 17

2.3.3.Implementation of Node2Vec 17

**3 RESULTS AND DISCUSSION** 19 3.1. Comparison of Existing Solutions 19

3.2. Data Collection and Performance metrics 20

**4** **CONCLUSION** 21

6.1 Conclusion 21

**5** **REFERENCES** 22

CMRCET B. Tech (CSE) Page No i

Link Prediction in Social Media network using machine learning approaches

**List of Figures**

**FIGURE**

**NO. TITLE PAGE NO.**

2.1.1 Architecture of encoder and decoder 3

2.1.2 Example of variational auto encoder 4

2.1.3 Architecture of the variational auto encoder 5

2.1.4 Adjacency matrix 6

2.1.5 Feature matrix 7

2.1.6 Architecture of Variational Graph Auto-encoder 8

2.2.1 Stochastic gradient descent 9

2.2.2 Regression Line 11

2.2.3 Least Square Regression Line 11

2.3.1 Generating n-dimensional node embeddings 13

from a input graph G using node2vec

2.3.2 Random walk 15

2.3.3 Example of generating training data for skip-gram model 15

2.3.4 Skip-Gram Model architecture 16

2.3.5 Node2Vec architecture 18

2.3.6 Block diagram of Node2vec architecture 18

3.1. Diagram of link prediction results of each algorithm 19

on 4 data sets

3.2 Prediction results of each algorithm on 4 datasets 20

CMRCET B. Tech (CSE) Page No ii

Link Prediction in Social Media network using machine learning approaches

# **ABSTRACT**

In Node2Vec, the global structure of the network is neglected and the stochastic gradient descent (SGD) method is easy to fall into local optimum. Based on this algorithm, an improved link prediction algorithm combining machine learning and hierarchical representation learning for network (HARP) is proposed. This method first uses adaptive learning optimizer Adam instead of SGD to improve Node2Vec, then divides the nodes and edges of the original network graph into a series of smaller layered graphs by merging them according to HARP, and then uses the improved Node2Vec algorithm to extract features continuously, so as to realize network embedding. Finally, a social network link prediction model based on machine learning and HARP is established. A series of social network link prediction experiments are carried out. The results show that the new method has excellent performance and stability.

CMRCET B. Tech (CSE) Page No iii

# **CHAPTER 1**

**INTRODUCTION**

Link Prediction in Social Media network using machine learning approaches

**CHAPTER 1**

**INTRODUCTION**

With the appearance of social media sites, the analysis of social networking services has received widespread attention in the past few years. The social network which consists of the relationships (edges) between persons, teams or communities (nodes) and nodes is ubiquitous and becomes the basic part of modern information infrastructure. Prediction of possible new links, as a basic element of graph mining, is applied in multiple fields. In social media, recommendation of friends and contents is often defined as an issue of link prediction . In addition, link prediction is applied to identify credit card fraud in the field of network security inspection, predict interaction between protein in biological information science, give shopping and film recommendation in e-commerce , and even identify hidden terrorist groups according to the activities of criminals and terrorists.

* 1. **PROBLEM STATEMENT**

Since a great number of social networks may contain millions of nodes and edges, it is extremely difficult to execute complicated reasoning processes in the whole network. A technology that has been used to deal with such matter is dimensionality reduction whose core idea is to search a mapping function and convert nodes in the graph into different latent representations with a lower dimension. Afterwards, these representations may be used as the characteristics of common tasks in the graph, such as classification of multiple tags, clustering and link prediction.

CMRCET B. Tech (CSE) Page No 1

Link Prediction in Social Media network using machine learning approaches

* 1. **RESEARCH OBJECTIVE**

With the appearance of social media sites, the analysis of social networking services has received widespread attention in the past few years.We will first divide the nodes and links of original network graph into a set of smaller layered graphs based on merging them according to HARP, a hierarchical representation learning algorithm proposed by Chen H et al. Then, we adopt the adaptive learning algorithm Adaptive Moment Estimation (Adam) to replace the optimizer SGD in Node2vec, utilize the improved Node2vec to extract features continuously, so as to realize final link prediction. Finally, we conduct link prediction in Twitter and Facebook datasets, and respectively compare the differences of similar link prediction algorithms of nodes and the methods of feature extraction with machine learning from the algorithm.

* 1. **PROJECT SCOPE AND LIMITATIONS**

As a future work we may involve other ML techniques and provide an extensive comparison among them. Furthermore, adding more software metrics in the learning process is one possible approach to increase the accuracy of the prediction model.

CMRCET B. Tech (CSE) Page No 2

**CHAPTER 2**

**BACKGROUND WORK**

Link Prediction in Social Media network using machine learning approaches

**CHAPTER 2**

**BACKGROUND WORK**

**2.1. VARIATIONAL GRAPH AUTO – ENCODER**

**2.1.1. INTODUCTION**

The traditional autoencoder is a neural network that contains an encoder and a decoder. The encoder takes a data point X as input and converts it to a lower-dimensional representation (embedding) Z. The decoder takes the lower-dimensional representation Z and returns a reconstruction of the original input X-hat that looks like the input X. The quality of the embedding determines the quality of the output X-hat. However, it might not be possible for the encoder to encode all information because the embedding has a lower dimensionality than the input. Therefore, if the embedding captures more information from the input, the output will have a better performance.

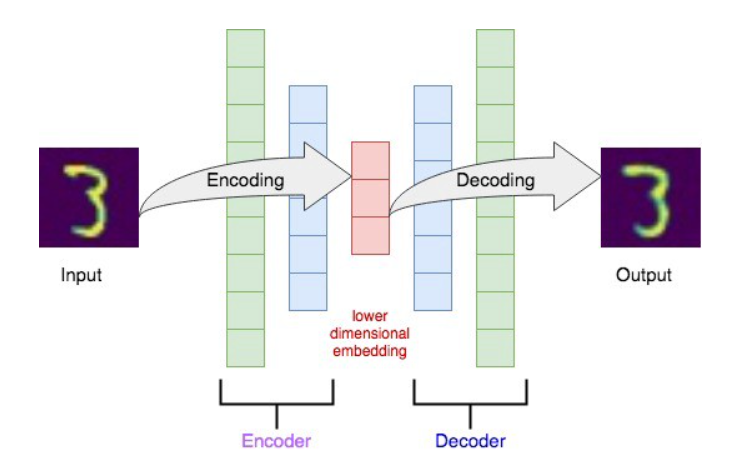
****

Fig 2.1.1. Architecture of the Encoder and Decoder

CMRCET B. Tech (CSE) Page No 3

Link Prediction in Social Media network using machine learning approaches

**Variational Auto-Encoder**

**Why do we need the variational autoencoders?**

One of the biggest advantages of the variational autoencoder is that VAE could generate new data from the original source dataset. In contrast, traditional autoencoder could only generate images that are similar to the original inputs.

Main idea

The main idea of a variational autoencoder is that it embeds the input X to a distribution rather than a point. And then a random sample Z is taken from the distribution rather than generated from encoder directly.

The encoder for a VAE is often written as qφ(z|x), which takes a data point X and produces a distribution. The distribution is usually parameterized as a multivariate Gaussian. Therefore, the encoder predicts the means and standard deviation of the Gaussian distribution. The lower-dimensional embedding Z is sampled from this distribution. The decoder is a variational approximation, pθ(x|z), which takes an embedding Z and produces the output X-hat.

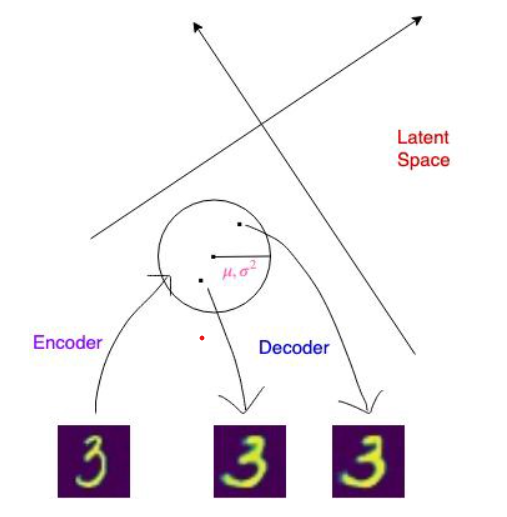


Fig 2.1.2. Example of variational auto encoder

CMRCET B. Tech (CSE) Page No 4

Link Prediction in Social Media network using machine learning approaches

The idea of VAE can be generalized by the image below:

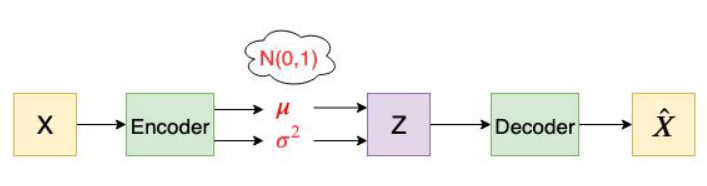


Fig 2.1.3. Architecture of the variational autoencoder

The encoder takes a datapoint X as input and generates μ and logσ² as outputs. The reason that we use logσ² instead of σ² is that σ² is non-negative, so we will need an extra activation function. But logσ² could be positive or negative. After we get μ and logσ², we try to make both μ and logσ² close to 0, which means μ is close to 0 and σ is close to 1. Therefore the final distribution will be close to N(0,1). Finally, we want to generate the embedding Z from μ and σ by z = μ + σ \* ε, where ε ~ N(0,1). This is called the reparameterization trick. Now with the latent variable Z, we can generate our output X-hat through the decoder.

**2.1.2. MERITS:**

An advantage for VAEs (Variational AutoEncoders) is that there is a clear and recognized way to evaluate the quality of the model (log-likelihood, either estimated by importance sampling or lower-bounded). Right now it’s not clear how to compare two GANs (Generative Adversarial Networks) or compare a GAN and other generative models except by visualizing samples.

**DEMERITS:**

A disadvantage of VAEs is that, because of the injected noise and imperfect reconstruction, and with the standard decoder (with factorized output distribution), the generated samples are much more blurred than those coming from GANs.

CMRCET B. Tech (CSE) Page No 5

Link Prediction in Social Media network using machine learning approaches

**2.1.3. Variational graph Auto - Encoder**

To build a variational graph autoencoder that applies the idea of VAE to graph-structured data. We want our variational graph autoencoder to be able to generate new graphs or reason about graphs. However, we can’t just straightforwardly apply the idea of VAE because graph-structured data are irregular. Each graph has a variable size of unordered nodes and each node in a graph has a different number of neighbours, so we can’t just use convolution directly anymore.

**Adjacency Matrix**

We use an adjacency matrix A to represent the input graph. Normally we assume the adjacency matrix is binary. The value 1 at row i and column j means that there is an edge between vertex i and vertex j. The value 0 at row m and column n means that there is no edge between vertex m and vertex n.

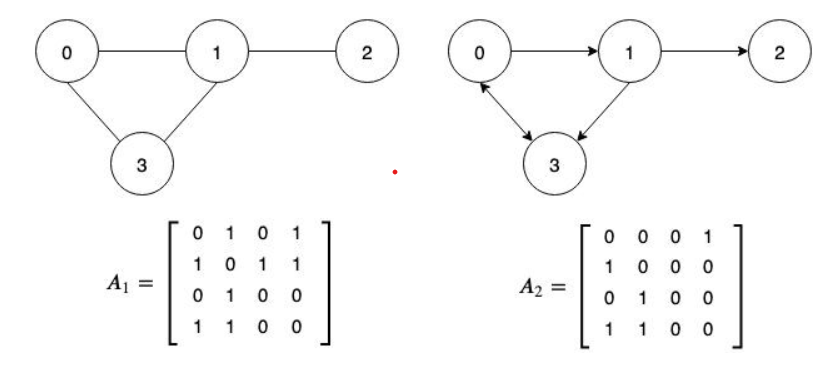
****

Fig 2.1.4. Adjacency matrix

CMRCET B. Tech (CSE) Page No 6

Link Prediction in Social Media network using machine learning approaches

**Feature Matrix**

We use the feature matrix X to represent the features of each node from the input graph. Row i of the feature matrix X represents the feature embeddings for vertex i.

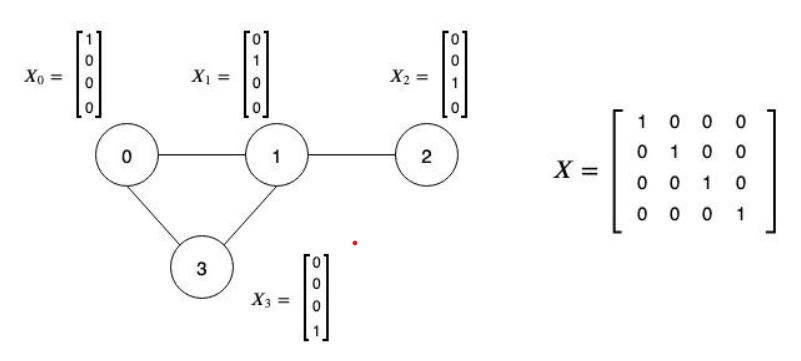
****

Fig 2.1.5. Feature matrix

CMRCET B. Tech (CSE) Page No 7 Link Prediction in Social Media network using machine learning approaches

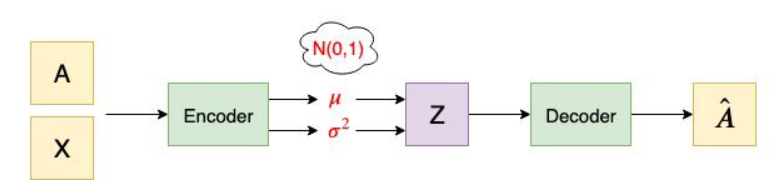


Fig 2.1.6. The architecture of variational graph autoencoder

The encoder (inference model) of VGAE consists of graph convolution networks (GCNs). It takes an adjacency matrix A and a feature matrix X as inputs and generates the latent variable Z as output. The first GCN layer generates a lower-dimensional feature matrix. Then we can calculate *Z* using parameterization trick

**Z =**

**2.1.2. MERITS:**

An advantage for VAEs (Variational AutoEncoders) is that there is a clear and recognized way to evaluate the quality of the model (log-likelihood, either estimated by importance sampling or lower-bounded). Right now it’s not clear how to compare two GANs (Generative Adversarial Networks) or compare a GAN and other generative models except by visualizing samples.

**DEMERITS:**

A disadvantage of VAEs is that, because of the injected noise and imperfect reconstruction, and with the standard decoder (with factorized output distribution), the generated samples are much more blurred than those coming from GANs.

CMRCET B. Tech (CSE) Page No 8

Link Prediction in Social Media network using machine learning approaches

**2.2. ADAPTIVE MOMENT ESTIMATION**

**2.2.1. INTRODUCTION**

Adam (Adaptive Moment Estimation) is a stochastic objective function optimization algorithm based on low-order moment adaptive estimation. This algorithm, in essence, is an RMSprop with the item of Momentum. It automatically modifies the learning rate of parameters by utilizing the first-order and the second-order moment evaluations of the gradient. The optimization idea of Adam is that after offset correction, any iterative learning rate is within a certain scope. It can bring more stable parameters than before. Combining Adagrad’s ability to deal with sparse gradients and RMSprop’s ability to cope with non-stationary targets, Adam has a weaker demand for memory and calculates different adaptive learning rates of different parameters. It can be used for great majority non-convex optimization objective functions, big datasets and high-dimensional spaces.

The optimization idea of Adam is that after offset correction, any iterative learning rate is within a certain scope. It can bring more stable parameters than before.

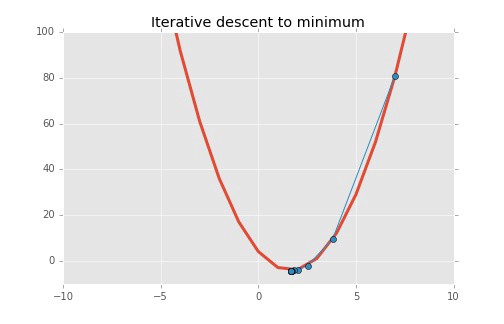


Fig 2.2.1. stochastic gradient descent

CMRCET B. Tech (CSE) Page No 9

Link Prediction in Social Media network using machine learning approaches

**2.2.2. MERITS**

1. Can handle sparse gradients on noisy datasets.
2. Default hyperparameter values do well on most problems.
3. Computationally efficient.
4. Requires little memory, thus memory efficient.

## DEMERITS

1. Adam does not converge to an optimal solution in some areas
2. Adam can suffer a weight decay problem (which is addressed in AdamW).
3. Recent optimization algorithms have been proven faster and better

**2.2.3 IMPLEMENTATION**

**Regression line :**

A regression line indicates a linear relationship between the dependent variables on the y-axis and the independent variables on the x-axis. The correlation is established by analyzing the data pattern formed by the variables.

CMRCET B. Tech (CSE) Page No 10

Link Prediction in Social Media network using machine learning approaches

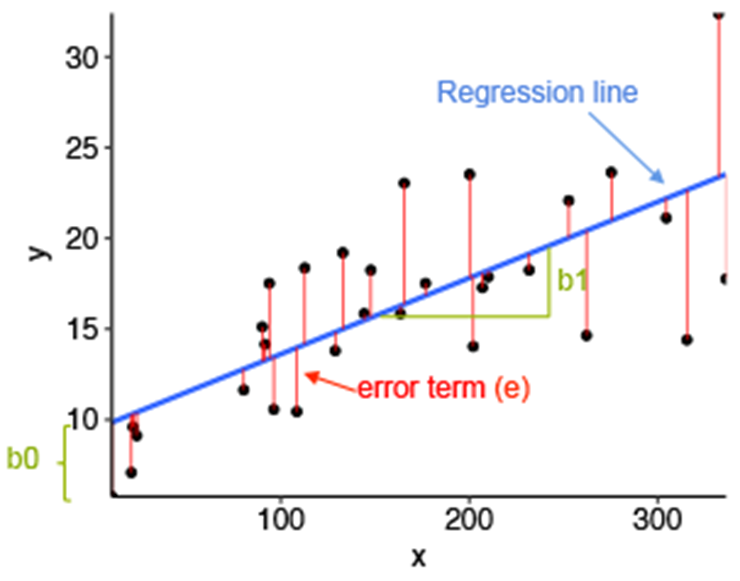


Fig 2.2.2. Regression Line

The Least Squares Regression Line (LSRL) is plotted nearest to the data points (x, y) on a regression graph.

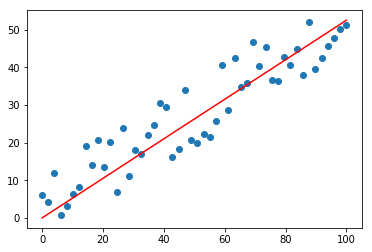


Fig 2.2.3. LSRL

CMRCET B. Tech (CSE) Page No 11

Link Prediction in Social Media network using machine learning approaches

Regression is widely used in financial models like CAPM and investing measures like Beta to determine the feasibility of a project. It is also used for creating projections of investments and financial returns.

If Y is the dependent variable and X is the independent variable, the Y on X regression line equation is represented as follows:

**‘Y = a + bX + €’**

Here,

* Y is the dependent variable.
* a is the Y-intercept.
* b is the slope of the regression line.
* X is the independent variable.
* ɛ is the residual (error).

Where SGD adds randomness to gradient descent, momentum accelerates the convergence, and adaptive gradient, as its name sounds, adapts to different learning rate for different parameters.

**Advantages:**

1. Can handle sparse gradients on noisy datasets.
2. Default hyperparameter values do well on most problems.
3. Computationally efficient.
4. Requires little memory, thus memory efficient.

## Disadvantages:

1. Adam does not converge to an optimal solution in some areas
2. Adam can suffer a weight decay problem (which is addressed in AdamW).
3. Recent optimization algorithms have been proven faster and better

CMRCET B. Tech (CSE) Page No 12

Link Prediction in Social Media network using machine learning approaches

**2.3. Node2Vec**

**2.3.1. INTRODUCTION**

**What is Node2Vec?**

A notable problem when working with networks, is transforming the network structure into numerical representation which can then be passed onto traditional machine learning algorithms. Node2Vec is an algorithm that allows the user to map nodes in a graph G to an embedding space. Generally, the embedding space is of lower dimensions than the number of nodes in the original graph G. The algorithm tries to preserve the initial structure within the original graph. Essentially, the nodes which are similar within the graph will yield similar embeddings in the embedding space. These embedding spaces are essentially a vector corresponding to each node in the network. The graph embeddings are commonly used as input features to solve machine learning problems oriented around link prediction community detection, classification, etc.

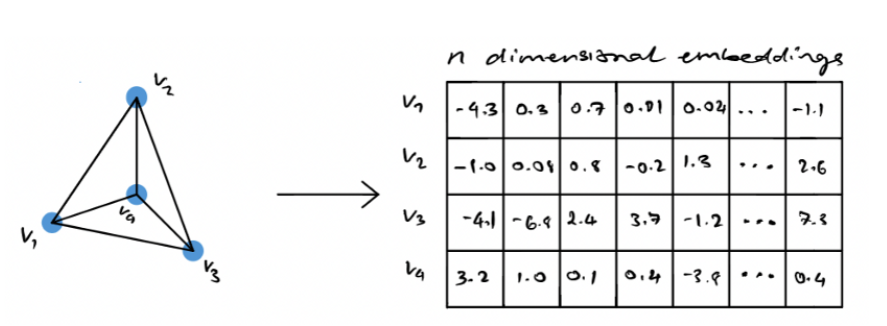


Fig 2.3.1. Generating n-dimensional node embeddings from a input graph G using node2vec

Generally, when dealing with very large graphs it’s quite difficult for scientists to visually represent the data they’re working with. A common solution to see how a graph looks is to generate node embeddings associated with that graph and then visualize the embeddings in a lower-dimensional space. This allows you to visually see potential clusters or groups forming in very large networks.

CMRCET B. Tech (CSE) Page No 13

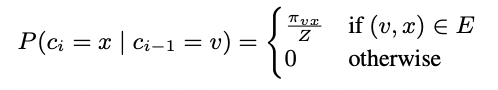
Link Prediction in Social Media network using machine learning approaches

**Random Walks Generation**

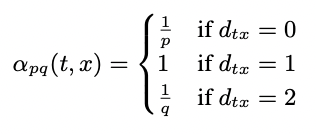
Having an understanding of what random walks are and how they work is crucial in understanding how node2vec works.

As a high level overview, the simplest comparison of a random walk would be through walking. Imagine that each step you take is determined probabilistically. This implies that at each index of time, you have moved in a certain direction based on a probabilistic outcome. This algorithm explores the relationship to each step that you would take and its distance from the initial starting point.

Now you might wonder how these probabilities of moving from one node to another are calculated. Node2Vec introduces the following formula for determining the probability of moving to the node x given that you were previously at the node v.



The paper states that the easiest way to introduce a bias to influence the random walks would be if there was a weight associated with each edge. However, that wouldn’t work in the case of unweighted networks. To resolve this, the authors introduced a guided random walk governed by two parameters p and q. p indicates the probability of a random walk getting back to the previous node, and q indicates the probability that a random walk can pass through a previously unseen part of the graph



CMRCET B. Tech (CSE) Page No 14

Link Prediction in Social Media network using machine learning approaches

Where dtx represents the shortest path between nodes t and x. It can be visually seen in the illustration below.

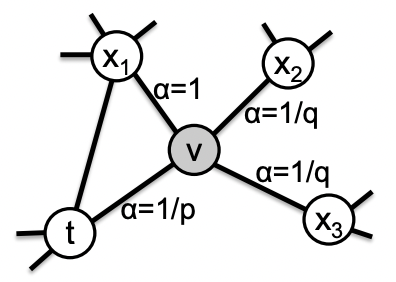


Fig 2.3.2. Random walk

**2.3.2. Skip-Gram Architecture**

The skip-gram model is a simple neural network with one hidden layer trained in order to predict the probability of a given word being present when an input word is present. The process can be described visually as seen below.

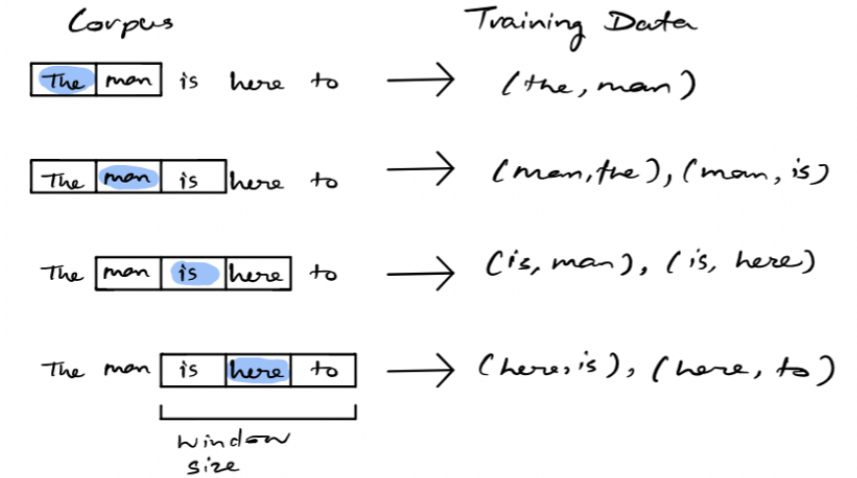
****

Fig 2.3.3. example of generating training data for skip-gram model

CMRCET B. Tech (CSE) Page No 15

Link Prediction in Social Media network using machine learning approaches

Given some corpus of text, a target word is selected over some rolling window. The training data consists of pairwise combinations of that target word and all other words in the window. This is the resulting training data for the neural network. Once the model is trained, we can essentially yield a probability of a word being a context word for a given target. The following image below represents the architecture of the neural network for the skip-gram model.

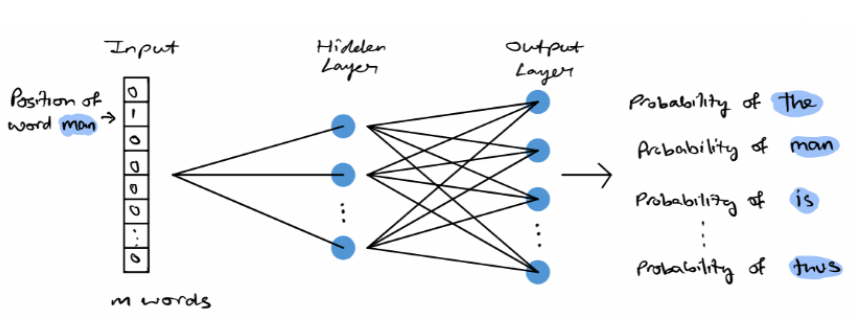


Fig 2.3.4. Skip-Gram Model architecture

A corpus can be represented as a vector of size N, where each element in N corresponds to a word in the corpus. During the training process, we have a pair of target and context words, the input array will have 0 in all elements except for the target word. The target word will be equal to 1. The hidden layer will learn the embedding representation of each word, yielding a d-dimensional embedding space [2]. The output layer is a dense layer with a softmax activation function. The output layer will essentially yield a vector of the same size as the input, each element in the vector will consist of a probability. This probability indicates the similarity between the target word and the associated word in the corpus.

CMRCET B. Tech (CSE) Page No 16

Link Prediction in Social Media network using machine learning approaches

**2.3.2. MERITS AND DEMERITS**

1. It’s scalable and parallelizes easily
2. Open sourced in python & spark
3. Unique approach to learning feature representation via node embeddings
4. The structure of the original network is preserved through the embeddings
5. Node2Vec has many real world applications including and not limited to node classification, community detection, link prediction, etc.

**2.3.3. IMPLEMTATION**

The process for node2vec is fairly simple, it begins by inputting a graph and extracting a set of random walks from the input graph. The walks can then be represented as a directed sequence of words where each node represents a word. The generated random walks are then passed into the skip-gram model. As explained above, the skip-gram model works on words and sentences, each node in the random walk can be represented as a word and the entire walk can be represented as a sentence. The result of the skip-gram model yields an embedding for each node (or word in this analogy). The entire process can be seen below.

CMRCET B. Tech (CSE) Page No 17

Link Prediction in Social Media network using machine learning approaches

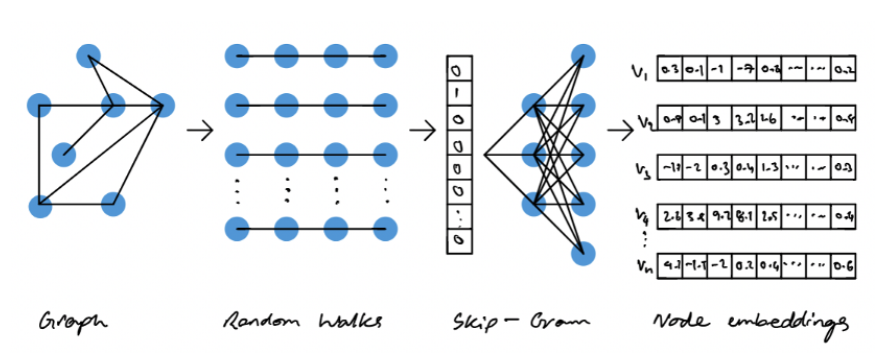


Fig 2.3.5. Node2Vec Architecture

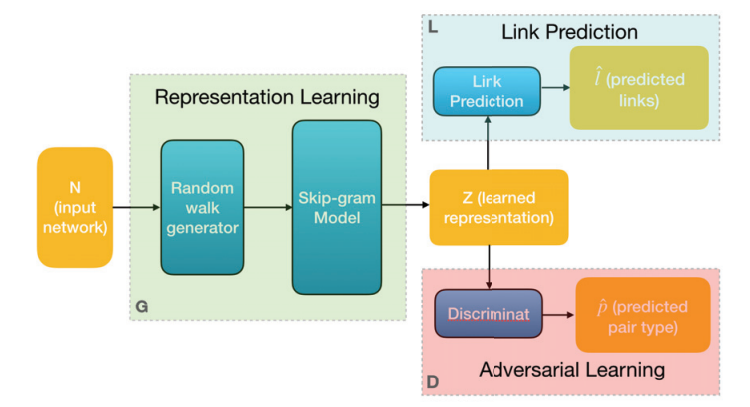


Fig 2.3.6. block diagram of Node2Vec

CMRCET B. Tech (CSE) Page No 18

**CHAPTER 3**

**RESULTS AND DISCUSSION**

Link Prediction in Social Media network using machine learning approaches

**CHAPTER 3**

**RESULTS AND DISCUSSION**

**3.1. Comparison of Existing Solutions**

Given the Node2Vec is a local approach which is limited to the structure around the node, it uses short random walk to find the local neighborhood of nodes. Such attention to local structure implicitly overlooks the long-distance relationship in whole network, and the studied representation may be unable to reveal the important global structure model. Meanwhile, the ‘meta-strategy’ HARP proposed by Chen et al. is a new paradigm for learning low-dimensional embedding of graph nodes, being used to enhance varied random walk methods through graph preprocessing steps. The algorithm incorporates the global structure relationship of the network, which can compensate for the deficiency of Node2vec in this regard. Hence, it names the algorithm HARP that combines the improved Node2vec and HARP(NewN2V).

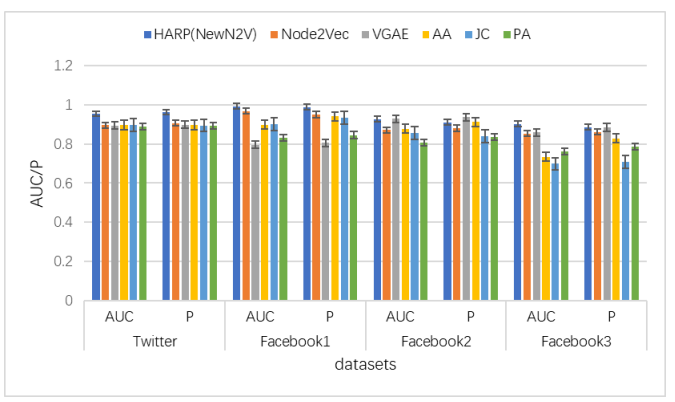


Fig 3.1. Diagram of link prediction results of each algorithm on 4 data sets

CMRCET B. Tech (CSE) Page No 19

Link Prediction in Social Media network using machine learning approaches

**3.2. Data Collection and Performance metrics**

The algorithm HARP (NewN2V) proposed was tested on the four datasets shown in table. In order to present the prediction results more accurately, the independent experiment was repeated on all datasets for 60 times, and the mean value of the AUC and precision (P) of these 60 experiments was calculated as the eventual outcome of link precision. Table provides the prediction results of each algorithm on 4 datasets. It vividly displays the prediction results of each algorithm on 4 datasets. Table also lists the degree of improvement in AUC indicators and precision of the proposed algorithms regarding 4 datasets compared with the Node2Vec algorithm

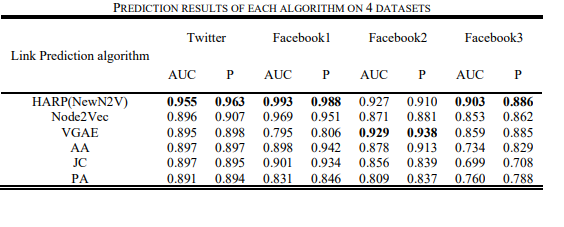


Fig 3.2. Prediction results of each algorithm on 4 datasets

CMRCET B. Tech (CSE) Page No 20

CHAPTER 4

**CONCLUSION**

Link Prediction in Social Media network using machine learning approaches

**CHAPTER 4**

**CONCLUSION**

It attempts to tackle the problems of Node2Vec algorithm in the process of optimizing non-convex functions, that is, the likeliness to run into the lousy local minimum value with unfavorable initialization due to lack of prior network information, and the inability of its optimizer SGD to conduct adaptive adjustment in the learning rate, which does not help with the treatment of sparse network. Besides, Node2Vec is aimed to retain local neighborhood of the graph, implicitly ignoring the global structure of the graph. Hence, this paper proposed to first use the machine learning optimizer Adam to replace SGD so as to optimize the Node2Vec algorithm. Meanwhile, HARP can merge input network graphs into smaller graphs yet with similar structures in a recursive manner to capture the global structure of input graphs, then study the graph representation of a set of smaller graphs and thus obtain a favorable initialization solution for input network graphs. Furthermore, such multi-level paradigm is combined with the improved Node2Vec algorithm to establish a link prediction model which is based on machine learning and HARP. The results of experiments conducted on real social network datasets show that the proposed algorithm is feasible and effective with certain superiority.

CMRCET B. Tech (CSE) Page No 21

**CHAPTER 5**

**REFERENCES**

Link Prediction in Social Media network using machine learning approaches

**CHAPTER 5**

## REFERENCES

## [1] D. Liben-Nowell, and J. Kleinberg, "The link-prediction problem for social networks", J Associat Inform Sci Tech, vol.58, no. (7), Jan 2007, pp. 1019-1031

## [2] L. M. Aiello, et al., "Friendship Prediction and Homophily in Social Media", Acm Transactions on the Web, vol.6, no. (2), May 2012, pp. 1-33

## [3] 3 P. Wang, X. U. Baowen, W. U. Yurong, and X. Y. Zhou, "Link prediction in social networks: the state-ofthe-art", Sci Chin, vol.58, no. (1), Jan 2015, pp. 11101- 011101

## [4] E. M. Airoldi, D. M. Blei, S. E. Fienberg, and E. P. Xing, "Mixed Membership Stochastic Blockmodels", J Mach Learn Res Jmlr, vol.9, no. (5), June 2008, pp. 1981

## [5] M. Bressan, et al., "Counting Graphlets: Space vs Time", Tenth Acm International Conference on Web Search & Data Mining, vol.6, no. (2), Feb 2017, pp. 06-10

## [6] H. C. Chen, B. Perozzi, Y. Hu, and S. Skiena, "HARP: Hierarchical Representation Learning for Networks, June 2017,

## [7] P. Pezeshkpour, Y. Tian, and S. Singh, "Investigating Robustness and Interpretability of Link Prediction via Adversarial Modifications, 2019,

## [8] B. Perozzi, R. Al-Rfou, and S. Skiena, "DeepWalk: Online Learning of Social Representations", Acm Sigkdd International Conference on Knowledge Discovery & Data Mining, Dec 2014, pp. 2-5

## [9] S. Cao, L. Wei, and Q. Xu: ‘GraRep: Learning Graph Representations with Global Structural Information’, in Editor (Ed.)^(Eds.): ‘Book GraRep: Learning Graph Representations with Global Structural Information’ (2015, edn.).

## [10] A. Grover, and J. Leskovec, "node2vec: Scalable Feature Learning for Networks", Acm Sigkdd International Conference on Knowledge Discovery & Data Mining, vol.1, no. (2), Aug 2016, pp. 855-864

CMRCET B. Tech (CSE) Page No 22